

#2

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/921,157

DATE: 08/15/2001

TIME: 09:26:47

Input Set : A:\Chir0315.app

Output Set: N:\CRF3\08132001\I921157.raw

ENTERED

3 <110> APPLICANT: Covacci, Antonio
 4 Bugnoli, Massimo
 5 Telford, John
 6 Macchia, Giovanni
 7 Rappuoli, Rino
 9 <120> TITLE OF INVENTION: Helicobacter Pylori Cytotoxin Proteins Useful For
 10 Vaccines And Diagnostics
 12 <130> FILE REFERENCE: CHIR0315
 C--> 14 <140> CURRENT APPLICATION NUMBER: US/09/921,157
 C--> 15 <141> CURRENT FILING DATE: 2001-08-02
 17 <150> PRIOR APPLICATION NUMBER: 08/466,662
 18 <151> PRIOR FILING DATE: 1995-06-06
 20 <150> PRIOR APPLICATION NUMBER: 08/256,848
 21 <151> PRIOR FILING DATE: 1994-10-21
 23 <150> PRIOR APPLICATION NUMBER: 09/360,934
 24 <151> PRIOR FILING DATE: 1999-07-26
 26 <150> PRIOR APPLICATION NUMBER: 08/471,491
 27 <151> PRIOR FILING DATE: 1995-06-06
 29 <150> PRIOR APPLICATION NUMBER: PCT/EP93/00472
 30 <151> PRIOR FILING DATE: 1993-03-02
 32 <150> PRIOR APPLICATION NUMBER: PCT/EP93/00158
 33 <151> PRIOR FILING DATE: 1993-01-25
 35 <150> PRIOR APPLICATION NUMBER: FI 92 A 000052
 36 <151> PRIOR FILING DATE: 1992-03-02
 38 <160> NUMBER OF SEQ ID NOS: 10
 40 <170> SOFTWARE: PatentIn Ver. 2.1
 42 <210> SEQ ID NO: 1
 43 <211> LENGTH: 27
 44 <212> TYPE: DNA
 45 <213> ORGANISM: Artificial Sequence ✓
 47 <220> FEATURE:
 48 <223> OTHER INFORMATION: Description of Artificial Sequence: Novel ✓
 49 Sequence
 51 <400> SEQUENCE: 1
 52 gcaagcttat cgatgtcgac tcgagct 27
 55 <210> SEQ ID NO: 2
 56 <211> LENGTH: 3960
 57 <212> TYPE: DNA
 58 <213> ORGANISM: Helicobacter pylori
 60 <400> SEQUENCE: 2
 61 aaaaagaaag gaagaaaatg gaaatacaac aaacacaccg caaaatcaat cgccctctgg 60
 62 tttctctcgc tttagtagga gcattagtca gcatcacacc gcaacaaagt catgccgcct 120
 63 ttttcacaac cgtgatcatt ccagccattg ttgggggtat cgctacaggc accgctgtag 180
 64 gaacgggtctc agggcttctt agctgggggc tcaaacaagc cgaagaagcc aataaaaccc 240
 65 cagataaacc cgataaagtt tggcgcattc aagcaggaaa aggctttaat gaattcccta 300
 66 acaaggaata cgacttatac agatcccttt tatccagtaa gattgatgga ggttgggatt 360
 67 gggggaatgc cgctaggcat tattgggtca aaggcgggca acagaataag cttgaagtgg 420

RAW SEQUENCE LISTING

DATE: 08/15/2001

PATENT APPLICATION: US/09/921,157

TIME: 09:26:47

Input Set : A:\Chir0315.app

Output Set: N:\CRF3\08132001\I921157.raw

```

68 atatgaaaga cgctgtaggg acttatacct tatcagggct tagaaacttt actggtgggg 480
69 atttagatgt caatatgcaa aaagccactt tacgcttggg ccaattcaat ggcaattctt 540
70 ttacaagcta taaggatagt gctgatcgca ccacgagagt ggatttcaac gctaaaaata 600
71 tctcaattga taattttgta gaaatcaaca atcgtgtggg ttctggagcc gggaggaaaag 660
72 ccagctctac ggttttgact ttgcaagctt cagaagggat cactagcgat aaaaacgctg 720
73 aaattttctt ttatgatggg gccacgctca atttggcttc aagcagcggt aaattaatgg 780
74 gtaatgtgtg gatgggcccgt ttgcaatacg tgggagcgta tttggcccct tcatacagca 840
75 cgataaacac ttcaaaagta acaggggaag tgaattttta ccacctcact gttggcgata 900
76 aaaacgccgc tcaagcgggc attatcgcta ataaaaagac taatattggc acactggatt 960
77 tgtggcaaaag cgccgggtta aacattatcg ctctccaga aggtggctat aaggataaac 1020
78 ccaataatac cccttctcaa agtgggtgcta aaaacgacaa aaatgaaagc gctaaaaacg 1080
79 acaaaacaaga gagcagtcaa aataatagta acactcaggt cattaaccca cccaatagtg 1140
80 cgcaaaaaac agaagttcaa ccacgcaag tcattgatgg gccttttgcg ggcgcaaaag 1200
81 acacggttgt caatatcaac cgcatcaaca ctaacgctga tggcacgatt agagtgggag 1260
82 ggttttaaagc ttctcttacc accaatgctg ctcatattgca tatcggcaaa ggcggtgtca 1320
83 atctgtccaa tcaagcgagc gggcgctctc ttatagtga aaatctaact gggaatatca 1380
84 ccgttgatgg gcctttaaga gtgaataatc aagtgggtgg ctatgctttg gcaggatcaa 1440
85 gcgcgaattt tgagtttaag gctggtacgg ataccaaaaa cggcacagcc acttttaata 1500
86 acgatattag tctgggaaga tttgtgaatt taaagtgga tgctcataca gctaatttta 1560
87 aaggtattga tacgggtaat ggtggtttca acacctaga ttttagtggc gttacagaca 1620
88 aagtcaatat caacaagctc attacggctt ccactaatgt ggccgttaaa aacttcaaca 1680
89 ttaatgaatt gattgttaaa accaatggga taagtgtggg ggaatatact cattttagcg 1740
90 aagatatagg cagtcaatcg cgcatcaata ccgtgcgttt ggaaactggc actagggtcac 1800
91 ttttctctgg ggggtttaa tttaaagggt gcgaaaaaatt ggttatagat gagttttact 1860
92 atagcccttg gaattatttt gacgctagaa atattaaaaa tgttgaaatc accaataaac 1920
93 ttgcttttgg acctcaagga agtccctggg gcacatcaaa acttatgttc aataatctaa 1980
94 ccctaggta aaatgcggtc atggattata gccaattttc aaatttaacc attcaagggg 2040
95 atttcatcaa caatcaaggc actatcaact atctggtccg aggtgggaaa gtggcaacct 2100
96 taagcgtagg caatgcagca gctatgatgt ttaataatga tatagacagc gcgaccggat 2160
97 tttacaaacc gctcatcaag attaacagcg ctcaagatct cattaataat acagaacatg 2220
98 ttttattgaa agcgaaaatc attggttatg gtaatgtttc tacagggtacc aatggcatta 2280
99 gtaatgttaa tctagaagag caattcaaaag agcgcctagc cttttataac aacaataacc 2340
100 gcatggatac ttgtgtggtg cgaaatactg atgacattaa agcatgcggt atggctatcg 2400
101 gcgatcaaag catggtgaac aaccctgaca attacaagta tcttatcggt aaggcatgga 2460
102 aaaatatagg gatcagcaaa acagctaatt gctctaaaat ttcggtgtat tatttaggca 2520
103 attctacgcc tactgagaat ggtggcaata ccacaaattt acccacaac accactagca 2580
104 atgcacgttc tgccaacaac gcccttgca aaaacgctcc tttcgctcaa cctagtgtca 2640
105 ctctaattt agtcgctatc aatcagcatg attttggcac tattgaaagc gtgtttgaat 2700
106 tggctaaccg ctctaaagat attgacacgc tttatgctaa ctcaggcgct caaggcaggg 2760
107 atctcttaca aaccttattg attgatagcc atgatgcggg ttatgccaga aaaatgattg 2820
108 atgctacaag cgctaataaa atcaccaagc aattgaatac ggccactacc actttaaaca 2880
109 acatagccag tttagagcat aaaaccagcg gcttacaac tttgagcttg agtaatgcga 2940
110 tgattttaaa ttctcgttta gtcaatctct ccaggagaca caccaaccat attgactcgt 3000
111 tcgccaacag cttacaagct ttaaaagacc aaaaattcgc ttctttagaa agcgcggcag 3060
112 aagtgttgta tcaatttgcc cctaaatatg aaaaacctac caatgtttgg gctaacgcta 3120
113 ttgggggaac gagcttgaat aatggctcta acgcttcatt gtatggcaca agcgcgggcg 3180
114 tagacgctta ccttaacggg caagtggaag ccattgtggg cggttttggg agctatgggt 3240
115 atagctcttt taataatcgt gcgaactccc ttaactctgg ggccaataac actaattttg 3300
116 gcgtgtatag ccgtattttt gcccaaccagc atgaatttga ctttgaagct caaggggcac 3360

```

RAW SEQUENCE LISTING

DATE: 08/15/2001

PATENT APPLICATION: US/09/921,157

TIME: 09:26:47

Input Set : A:\Chir0315.app

Output Set: N:\CRF3\08132001\I921157.raw

```

117 tagggagcga tcaatcaagc ttgaatttca aaagcgctct attacaagat ttgaatcaaa 3420
118 gctatcatta cttagcctat agcgctgcaa caagagcgag ctatgggttat gacttcgcgt 3480
119 tttttaggaa cgcttttagtg ttaaaaccaa gcgtgggtgt gagctataac catttaggtt 3540
120 caaccaactt taaaagcaac agcaccaatc aagtggcttt gaaaaatggc tctagcagtc 3600
121 agcatttatt caacgctagc gctaattgtg aagcgcgcta ttattatggg gacacttcat 3660
122 acttctacat gaatgctgga gttttacaag agttcgctca tgttggctct aataacgccg 3720
123 cgtctttaaa cacctttaaa gtgaatgccg ctcgcaaccc tttaaatacc catgccagag 3780
124 tgatgatggg tggggaatta aaattagcta aagaagtgtt tttgaatttg ggcgttggtt 3840
125 atttgcacaa tttgatttcc aatataggcc atttcgcttc caatttagga atgagggtata 3900
126 gtttctaaat accgctctta aacctatgct caaagcatgg gtttgaaatc ttacaaaaca 3960
129 <210> SEQ ID NO: 3
130 <211> LENGTH: 1296
131 <212> TYPE: PRT
132 <213> ORGANISM: Helicobacter pylori
134 <400> SEQUENCE: 3
135 Met Glu Ile Gln Gln Thr His Arg Lys Ile Asn Arg Pro Leu Val Ser
136 1 5 10 15
138 Leu Ala Leu Val Gly Ala Leu Val Ser Ile Thr Pro Gln Gln Ser His
139 20 25 30
141 Ala Ala Phe Phe Thr Thr Val Ile Ile Pro Ala Ile Val Gly Gly Ile
142 35 40 45
144 Ala Thr Gly Thr Ala Val Gly Thr Val Ser Gly Leu Leu Ser Trp Gly
145 50 55 60
147 Leu Lys Gln Ala Glu Glu Ala Asn Lys Thr Pro Asp Lys Pro Asp Lys
148 65 70 75 80
150 Val Trp Arg Ile Gln Ala Gly Lys Gly Phe Asn Glu Phe Pro Asn Lys
151 85 90 95
153 Glu Tyr Asp Leu Tyr Arg Ser Leu Leu Ser Ser Lys Ile Asp Gly Gly
154 100 105 110
156 Trp Asp Trp Gly Asn Ala Ala Arg His Tyr Trp Val Lys Gly Gly Gln
157 115 120 125
159 Gln Asn Lys Leu Glu Val Asp Met Lys Asp Ala Val Gly Thr Tyr Thr
160 130 135 140
162 Leu Ser Gly Leu Arg Asn Phe Thr Gly Gly Asp Leu Asp Val Asn Met
163 145 150 155 160
165 Gln Lys Ala Thr Leu Arg Leu Gly Gln Phe Asn Gly Asn Ser Phe Thr
166 165 170 175
168 Ser Tyr Lys Asp Ser Ala Asp Arg Thr Thr Arg Val Asp Phe Asn Ala
169 180 185 190
171 Lys Asn Ile Ser Ile Asp Asn Phe Val Glu Ile Asn Asn Arg Val Gly
172 195 200 205
174 Ser Gly Ala Gly Arg Lys Ala Ser Ser Thr Val Leu Thr Leu Gln Ala
175 210 215 220
177 Ser Glu Gly Ile Thr Ser Asp Lys Asn Ala Glu Ile Ser Leu Tyr Asp
178 225 230 235 240
180 Gly Ala Thr Leu Asn Leu Ala Ser Ser Ser Val Lys Leu Met Gly Asn
181 245 250 255
183 Val Trp Met Gly Arg Leu Gln Tyr Val Gly Ala Tyr Leu Ala Pro Ser
184 260 265 270

```

RAW SEQUENCE LISTING

DATE: 08/15/2001

PATENT APPLICATION: US/09/921,157

TIME: 09:26:47

Input Set : A:\Chir0315.app

Output Set: N:\CRF3\08132001\I921157.raw

```

186 Tyr Ser Thr Ile Asn Thr Ser Lys Val Thr Gly Glu Val Asn Phe Asn
187      275      280      285
189 His Leu Thr Val Gly Asp Lys Asn Ala Ala Gln Ala Gly Ile Ile Ala
190      290      295      300
192 Asn Lys Lys Thr Asn Ile Gly Thr Leu Asp Leu Trp Gln Ser Ala Gly
193 305      310      315      320
195 Leu Asn Ile Ile Ala Pro Pro Glu Gly Gly Tyr Lys Asp Lys Pro Asn
196      325      330      335
198 Asn Thr Pro Ser Gln Ser Gly Ala Lys Asn Asp Lys Asn Glu Ser Ala
199      340      345      350
201 Lys Asn Asp Lys Gln Glu Ser Ser Gln Asn Asn Ser Asn Thr Gln Val
202      355      360      365
204 Ile Asn Pro Pro Asn Ser Ala Gln Lys Thr Glu Val Gln Pro Thr Gln
205      370      375      380
207 Val Ile Asp Gly Pro Phe Ala Gly Gly Lys Asp Thr Val Val Asn Ile
208 385      390      395      400
210 Asn Arg Ile Asn Thr Asn Ala Asp Gly Thr Ile Arg Val Gly Gly Phe
211      405      410      415
213 Lys Ala Ser Leu Thr Thr Asn Ala Ala His Leu His Ile Gly Lys Gly
214      420      425      430
216 Gly Val Asn Leu Ser Asn Gln Ala Ser Gly Arg Ser Leu Ile Val Glu
217      435      440      445
219 Asn Leu Thr Gly Asn Ile Thr Val Asp Gly Pro Leu Arg Val Asn Asn
220      450      455      460
222 Gln Val Gly Gly Tyr Ala Leu Ala Gly Ser Ser Ala Asn Phe Glu Phe
223 465      470      475      480
225 Lys Ala Gly Thr Asp Thr Lys Asn Gly Thr Ala Thr Phe Asn Asn Asp
226      485      490      495
228 Ile Ser Leu Gly Arg Phe Val Asn Leu Lys Val Asp Ala His Thr Ala
229      500      505      510
231 Asn Phe Lys Gly Ile Asp Thr Gly Asn Gly Gly Phe Asn Thr Leu Asp
232      515      520      525
234 Phe Ser Gly Val Thr Asp Lys Val Asn Ile Asn Lys Leu Ile Thr Ala
235      530      535      540
237 Ser Thr Asn Val Ala Val Lys Asn Phe Asn Ile Asn Glu Leu Ile Val
238 545      550      555      560
240 Lys Thr Asn Gly Ile Ser Val Gly Glu Tyr Thr His Phe Ser Glu Asp
241      565      570      575
243 Ile Gly Ser Gln Ser Arg Ile Asn Thr Val Arg Leu Glu Thr Gly Thr
244      580      585      590
246 Arg Ser Leu Phe Ser Gly Gly Val Lys Phe Lys Gly Gly Glu Lys Leu
247      595      600      605
249 Val Ile Asp Glu Phe Tyr Tyr Ser Pro Trp Asn Tyr Phe Asp Ala Arg
250      610      615      620
252 Asn Ile Lys Asn Val Glu Ile Thr Asn Lys Leu Ala Phe Gly Pro Gln
253 625      630      635      640
255 Gly Ser Pro Trp Gly Thr Ser Lys Leu Met Phe Asn Asn Leu Thr Leu
256      645      650      655
258 Gly Gln Asn Ala Val Met Asp Tyr Ser Gln Phe Ser Asn Leu Thr Ile

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/921,157

DATE: 08/15/2001

TIME: 09:26:47

Input Set : A:\Chir0315.app

Output Set: N:\CRF3\08132001\I921157.raw

259				660				665				670				
261	Gln	Gly	Asp	Phe	Ile	Asn	Asn	Gln	Gly	Thr	Ile	Asn	Tyr	Leu	Val	Arg
262				675				680				685				
264	Gly	Gly	Lys	Val	Ala	Thr	Leu	Ser	Val	Gly	Asn	Ala	Ala	Ala	Met	Met
265			690				695					700				
267	Phe	Asn	Asn	Asp	Ile	Asp	Ser	Ala	Thr	Gly	Phe	Tyr	Lys	Pro	Leu	Ile
268	705					710					715				720	
270	Lys	Ile	Asn	Ser	Ala	Gln	Asp	Leu	Ile	Lys	Asn	Thr	Glu	His	Val	Leu
271					725					730					735	
273	Leu	Lys	Ala	Lys	Ile	Ile	Gly	Tyr	Gly	Asn	Val	Ser	Thr	Gly	Thr	Asn
274				740					745					750		
276	Gly	Ile	Ser	Asn	Val	Asn	Leu	Glu	Glu	Gln	Phe	Lys	Glu	Arg	Leu	Ala
277				755				760					765			
279	Leu	Tyr	Asn	Asn	Asn	Asn	Arg	Met	Asp	Thr	Cys	Val	Val	Arg	Asn	Thr
280		770					775					780				
282	Asp	Asp	Ile	Lys	Ala	Cys	Gly	Met	Ala	Ile	Gly	Asp	Gln	Ser	Met	Val
283	785					790					795				800	
285	Asn	Asn	Pro	Asp	Asn	Tyr	Lys	Tyr	Leu	Ile	Gly	Lys	Ala	Trp	Lys	Asn
286					805						810				815	
288	Ile	Gly	Ile	Ser	Lys	Thr	Ala	Asn	Gly	Ser	Lys	Ile	Ser	Val	Tyr	Tyr
289				820					825					830		
291	Leu	Gly	Asn	Ser	Thr	Pro	Thr	Glu	Asn	Gly	Gly	Asn	Thr	Thr	Asn	Leu
292			835					840					845			
294	Pro	Thr	Asn	Thr	Thr	Ser	Asn	Ala	Arg	Ser	Ala	Asn	Asn	Ala	Leu	Ala
295		850					855					860				
297	Gln	Asn	Ala	Pro	Phe	Ala	Gln	Pro	Ser	Ala	Thr	Pro	Asn	Leu	Val	Ala
298	865				870						875				880	
300	Ile	Asn	Gln	His	Asp	Phe	Gly	Thr	Ile	Glu	Ser	Val	Phe	Glu	Leu	Ala
301					885					890				895		
303	Asn	Arg	Ser	Lys	Asp	Ile	Asp	Thr	Leu	Tyr	Ala	Asn	Ser	Gly	Ala	Gln
304				900					905				910			
306	Gly	Arg	Asp	Leu	Leu	Gln	Thr	Leu	Leu	Ile	Asp	Ser	His	Asp	Ala	Gly
307			915					920					925			
309	Tyr	Ala	Arg	Lys	Met	Ile	Asp	Ala	Thr	Ser	Ala	Asn	Glu	Ile	Thr	Lys
310		930					935					940				
312	Gln	Leu	Asn	Thr	Ala	Thr	Thr	Thr	Leu	Asn	Asn	Ile	Ala	Ser	Leu	Glu
313	945				950					955					960	
315	His	Lys	Thr	Ser	Gly	Leu	Gln	Thr	Leu	Ser	Leu	Ser	Asn	Ala	Met	Ile
316				965						970				975		
318	Leu	Asn	Ser	Arg	Leu	Val	Asn	Leu	Ser	Arg	Arg	His	Thr	Asn	His	Ile
319				980					985				990			
321	Asp	Ser	Phe	Ala	Lys	Arg	Leu	Gln	Ala	Leu	Lys	Asp	Gln	Lys	Phe	Ala
322			995				1000					1005				
324	Ser	Leu	Glu	Ser	Ala	Ala	Glu	Val	Leu	Tyr	Gln	Phe	Ala	Pro	Lys	Tyr
325		1010				1015					1020					
327	Glu	Lys	Pro	Thr	Asn	Val	Trp	Ala	Asn	Ala	Ile	Gly	Gly	Thr	Ser	Leu
328	1025				1030					1035					1040	
330	Asn	Asn	Gly	Ser	Asn	Ala	Ser	Leu	Tyr	Gly	Thr	Ser	Ala	Gly	Val	Asp
331				1045					1050					1055		

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/921,157

DATE: 08/15/2001

TIME: 09:26:48

Input Set : A:\Chir0315.app

Output Set: N:\CRF3\08132001\I921157.raw

L:14 M:270 C: Current Application Number differs, Replaced Application Number

L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date